

EVALUASI SENTIMEN TERHADAP IMPLEMENTASSI CORETAX MENGGUNAKAN INDOBERTWEET DAN BERTOPIC

Selfy Ilafia¹, Daniel Arsa², Mutia fadhila Putri³

1,2,3) Sestem Informasi, Fakultas Sains dan Teknologi, Universitas jambi, Indonesia

Article Info	ABSTRACT
Article history: Received: 05 Juni 2026 Revised: 10 Juni 2026 Accepted: 18 Juni 2026	Abstrak Coretax (Sistem Administrasi Perpajakan Inti) merupakan sistem administrasi perpajakan yang diimplementasikan oleh Direktorat Jenderal Pajak (DJP) pada 1 Januari 2025. Implementasi sistem ini memunculkan beragam respons dari publik di media sosial, khususnya platform X. Penelitian ini bertujuan menganalisis sentimen dan pemodelan topik terhadap tweet terkait Coretax menggunakan IndoBERTweet dan BERTopic. Hasil evaluasi model IndoBERTweet menunjukkan akurasi terbaik sebesar 92% dengan F1-score kelas positif 69% pada pendekatan tanpa preprocessing tanpa normalisasi, stopword removal, dan stemming dibandingkan pendekatan dengan preprocessing yang menghasilkan akurasi 90% dan F1-score kelas positif 65%. BERTopic berhasil mengidentifikasi 14 topik utama dengan rata-rata coherence score 0,54, yang didominasi oleh isu teknis seperti error sistem, login, faktur pajak, NPWP, dan pelaporan SPT. Secara keseluruhan, sentimen publik terhadap implementasi Coretax didominasi sentimen negatif yang berpusat pada kendala teknis dan permasalahan operasional sistem. Kata Kunci: Coretax, IndoBERTweet, Analisis Sentimen, BERTopic, Pemodelan Topik Abstract Coretax (Core Tax Administration System) is a tax administration system implemented by the Directorate General of Taxes (DGT) on January 1, 2025. The rollout of this system generated diverse public responses on social media, particularly on the X platform. This study aims to analyze sentiment and topic modeling of Coretax-related tweets using IndoBERTweet and BERTopic. Evaluation results of the IndoBERTweet model show a best accuracy of 92% with a positive-class F1-score of 69% under the no-preprocessing approach without normalization, stopword removal, or stemming, compared to the preprocessing approach which includes normalization, stopword removal, and stemming yielding an accuracy of 90% and a positive-class F1-score of 65%. BERTopic successfully identified 14 main topics with an average coherence score of 0.54, dominated by technical issues such as system errors, login problems, tax invoices, NPWP registration, and SPT reporting. Overall, public sentiment toward the implementation of Coretax is predominantly negative, centering on technical obstacles and operational issues within the system. Keywords: Sentiment Analysis, BERTopic, Coretax, IndoBERTweet, Topic Modeling Djtechno: Jurnal Teknologi Informasi oleh Universitas Dharmawangsa Artikel ini bersifat open access yang didistribusikan di bawah syarat dan ketentuan dengan Lisensi Internasional Creative Commons Attribution NonCommercial ShareAlike 4.0 (CC-BY-NC-SA). 
Corresponding Author: E-mail : selfyilafia01@gmail.com	

1. PENDAHULUAN

Pajak merupakan sumber pendapatan utama pemerintah dan memainkan peran penting dalam mendukung pembangunan nasional. Optimalisasi sistem perpajakan menjadi prioritas pemerintah untuk meningkatkan kepatuhan wajib pajak, memperluas basis pajak, dan meminimalkan kebocoran pendapatan. Salah satu upaya strategis yang dilakukan adalah transformasi digital sistem administrasi perpajakan melalui pengembangan Sistem Administrasi Perpajakan Inti (Coretax) oleh Direktorat Jenderal Pajak (DJP).

Penerapan sistem administrasi ini bertujuan untuk meningkatkan efisiensi, transparansi, dan akuntabilitas administrasi perpajakan di Indonesia [1]. Sistem ini secara resmi diterapkan pada 1 Januari 2025, dengan nilai proyek sekitar Rp1,2 triliun. Namun, pada pertengahan 2025, penerimaan pajak aktual baru mencapai 53,8% dari target APBN, dengan tingkat pertumbuhan tahunan sebesar 3,1% [2]. Penerapan Coretax telah memicu beragam reaksi publik di media sosial, terutama di platform X. Sebagian masyarakat meyakini bahwa Coretax dapat meningkatkan efisiensi layanan administrasi perpajakan, sementara yang lain mengkritik berbagai tantangan teknis dan biaya implementasi yang tinggi. Oleh karena itu, analisis sentimen merupakan pendekatan yang relevan untuk mengidentifikasi tren opini publik terkait penerapan Coretax.

Analisis sentimen merupakan komponen dari *Natural Language Processing* (NLP) yang digunakan untuk mengidentifikasi dan mengklasifikasikan opini ke dalam kategori positif, negatif, atau netral [3]. Beberapa penelitian menunjukkan bahwa model berbasis Transformer memiliki performa yang unggul dalam tugas analisis sentimen berbahasa Indonesia. Penelitian oleh [4] menunjukkan bahwa model IndoBERTweet memperoleh kinerja yang lebih baik dibandingkan IndoBERT dalam analisis sentimen terkait kendaraan listrik di Indonesia, ditinjau dari nilai *accuracy*, *precision*, *recall*, dan *F1-score*. Selain itu, penelitian [5] juga menunjukkan bahwa pendekatan berbasis Transformer, khususnya IndoBERT dan IndoBERTweet, mampu memberikan performa klasifikasi sentimen yang lebih baik pada data media sosial dibandingkan pendekatan konvensional.

Pada pemodelan topik, metode BERTopic semakin banyak digunakan karena mampu memanfaatkan representasi semantik dari model *Transformer*. [5] berhasil menerapkan BERTopic untuk mengidentifikasi topik utama dalam pemberitaan Pemilu 2024 dan menghasilkan topik yang mudah diinterpretasikan. Selain itu, [6] menunjukkan

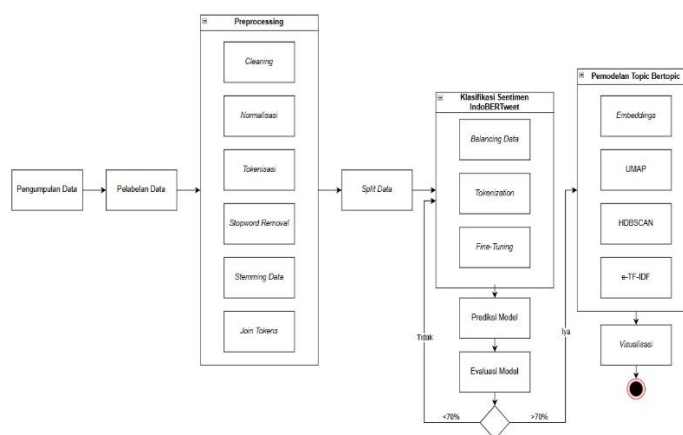
bahwa BERTopic memiliki nilai koherensi topik yang lebih tinggi dibandingkan metode *Latent Dirichlet Allocation* (LDA), sehingga lebih efektif dalam mengidentifikasi tren dan tema utama pada kumpulan dokumen teks.

Meskipun berbagai penelitian telah menunjukkan bahwa model berbasis *Transformer* efektif digunakan untuk analisis sentimen dan BERTopic mampu menghasilkan topik yang lebih representatif, penelitian yang membahas respons publik terhadap implementasi Coretax di Indonesia masih cukup terbatas. Selain itu, sebagian besar penelitian sebelumnya hanya berfokus pada analisis sentimen atau pemodelan topik secara terpisah, sehingga belum memberikan gambaran yang utuh mengenai sentimen masyarakat beserta isu-isu yang menjadi perhatian utama.

Analisis sentimen dan pemodelan topik terhadap Coretax tidak hanya penting untuk mengetahui persepsi masyarakat, tetapi juga untuk memperluas penerapan metode *Natural Language Processing* (NLP). Dengan menggabungkan analisis sentimen dan pemodelan topik dapat memberikan pemahaman yang lebih komprehensif mengenai hubungan antara opini masyarakat dan isu-isu yang mendasarinya.

Oleh karena itu, penelitian ini dilakukan untuk menganalisis sentimen masyarakat dan mengidentifikasi topik-topik yang banyak dibahas pengguna media sosial X terkait implementasi Coretax menggunakan metode IndoBERTweet dan BERTopic.

2. METODE PENELITIAN



Gambar 1. Tahapan Penelitian

2.1 Pengumpulan Data

Pengumpulan data dilakukan melalui proses scraping menggunakan *tools Tweet Harvest* yang berbasis *Playwright*. Data yang diambil berupa *tweet* yang mengandung

kata kunci "Coretax" dalam rentang waktu 1 Maret 2025 hingga 30 Juni 2025. Rentang waktu ini dipilih untuk melihat gambaran respons publik pada masa awal implementasi Coretax. Pada tahap ini, berbagai opini, pengalaman pengguna, dan berbagai kendala mulai banyak disampaikan melalui media sosial, sehingga dapat memberikan gambaran mengenai respons publik terhadap implementasi Coretax.

2.2 Pelabelan Data

Pelabelan data dilakukan secara manual oleh peneliti dan divalidasi dengan bantuan salah satu rekan ahli bahasa dari Universitas Jambi. Proses pelabelan dilakukan dengan mengklasifikasikan setiap *tweet* ke dalam dua kategori sentimen, yaitu positif dan negatif. Pemilihan dua kategori tersebut didasarkan pada karakteristik opini pengguna di media sosial yang cenderung ekspresif dan secara langsung menunjukkan kecenderungan sikap terhadap suatu isu.

2.3 Tahapan Preprocessing

Tahap preprocessing meliputi *cleaning*, *case folding*, normalisasi, tokenisasi, *stopword removal*, *stemming*, dan *join tokens*. Tahapan ini bertujuan membersihkan serta menyeragamkan teks agar siap digunakan pada proses analisis sentimen dan pemodelan topik [7]. Contoh hasil preprocessing ditunjukkan pada Tabel 1.

Tabel 1. Data Hasil *Preprocessing*

No	Data	Hasil <i>Preprocessing</i>
1	good job pemerintah! Coretax bisa bantu ekonomi makin top https://t.co/8mXm8SoXx6	good job pemerintah coretax bantu ekonomi makin top
2	yaudahhh bikin proyek coretax part II gitu aja koq repot! ■	Yaudah bikin proyek coretax part ii repot

2.4 Analisis Sentimen IndoBERTweet

Analisis sentimen pada penelitian ini dilakukan menggunakan model IndoBERTweet. Pemilihan IndoBERTweet didasarkan pada kemampuannya dalam memahami konteks bahasa informal pada media sosial, khususnya X, sehingga dinilai efektif untuk menganalisis opini masyarakat terkait implementasi Coretax [8].

IndoBERTweet merupakan model pre-trained Transformer berbasis *Bidirectional Encoder Representations from Transformers* (BERT) yang dikembangkan khusus untuk teks media sosial berbahasa Indonesia [9]. Model ini dilatih menggunakan data *tweet* berbahasa Indonesia sehingga mampu memahami karakteristik bahasa informal, singkatan, slang, dan pola bahasa yang umum digunakan pada media sosial.

2.5 Pemodelan Topik BERTopic

Pemodelan topik merupakan teknik dalam *Natural Language Processing* (NLP) yang digunakan untuk mengidentifikasi topik-topik utama dari sekumpulan teks secara otomatis. Pada penelitian ini, pemodelan topik dilakukan menggunakan BERTopic untuk mengetahui topik pembahasan dominan terkait implementasi Coretax [10].

BERTopic merupakan metode pemodelan topik berbasis *Transformer* yang memanfaatkan representasi semantik teks (*embedding*) untuk menghasilkan topik yang lebih kontekstual dan representatif. BERTopic menggabungkan beberapa tahapan, yaitu *sentence embedding*, reduksi dimensi, *clustering*, dan representasi topik [9]. Reduksi dimensi dilakukan menggunakan UMAP, sedangkan clustering menggunakan HDBSCAN. Representasi topik dihasilkan menggunakan *class-Based Term Frequency Inverse Document Frequency* (c-TF-IDF) untuk menghitung bobot kata berdasarkan distribusi dalam kluster tertentu dibandingkan dengan keseluruhan korpus [11].

2.6 Evaluasi Model

Evaluasi model dilakukan untuk mengukur performa analisis sentimen IndoBERTweet dan kualitas topik yang dihasilkan oleh BERTopic. Evaluasi analisis sentimen dilakukan menggunakan confusion matrix dan metrik *accuracy*, *precision*, *recall*, dan *F1-score* pada data *testing* [12].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Evaluasi pemodelan topik dilakukan menggunakan coherence score untuk mengukur keterkaitan semantik antar kata dalam suatu topik. Semakin tinggi nilai coherence score, maka semakin baik kualitas dan interpretabilitas topik yang dihasilkan [8].

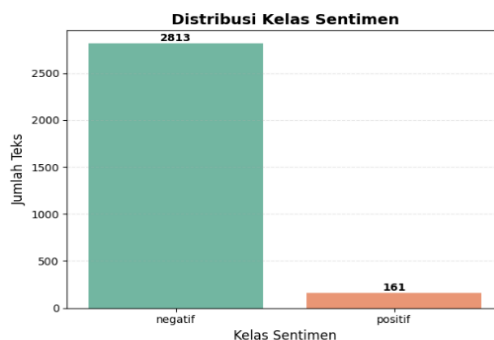
$$Coherence\ Score = \frac{1}{N} \sum_{i=1}^N score(w_i) \quad (5)$$

Dengan N menyatakan jumlah kata dalam topik dan $score(w_i)$ merupakan nilai keterkaitan semantik dari setiap kata pada topik tersebut.

3. HASIL DAN PEMBAHASAN

3.1 Pengambilan dan Pelabelan Data

Pengumpulan data dilakukan menggunakan metode *scraping* dari media sosial X dengan bantuan *Tweet Harvest*. Proses *scraping* menggunakan kata kunci "Coretax" pada periode 1 Maret 2025 hingga 30 Juni 2025. Dari proses tersebut diperoleh sebanyak 5.426 *tweet*. Setelah dilakukan pembersihan data untuk menghapus *tweet* yang tidak relevan dengan topik penelitian, jumlah data yang digunakan menjadi 2.974 *tweet*.



Gambar 2. Kelas Sentimen

Hasil pelabelan menunjukkan terdapat 2.813 data sentimen negatif dan 161 data sentimen positif. Distribusi tersebut menunjukkan adanya ketidakseimbangan data (*imbalanced data*) sehingga diperlukan proses penyeimbangan menggunakan *Random Oversampling*.

3.1.1 Pre-Processing

Tahap *preprocessing* dilakukan untuk membersihkan dan menyiapkan data teks sebelum proses analisis. Tahapan yang dilakukan meliputi *cleaning*, *case folding*, normalisasi, tokenisasi, *stopword removal*, *stemming*, dan *join tokens*.

Penelitian ini menggunakan dua pendekatan *preprocessing*, yaitu:

1. Menggunakan normalisasi, *stopword removal*, dan *stemming*.
2. Tanpa normalisasi, *stopword removal*, dan *stemming*.

Pendekatan pertama menghasilkan teks yang lebih bersih dan terstruktur, sedangkan pendekatan kedua mempertahankan konteks asli teks media sosial.

3.1.2 Pembagian Data

Dataset dibagi menjadi *training set*, *validation set*, dan *test set* dengan proporsi 70:15:15 menggunakan *stratified split* untuk mempertahankan distribusi kelas pada setiap dataset [8]. Hasil pembagian data disajikan pada Tabel 2 berikut.

Tabel 2. Pembagian Data

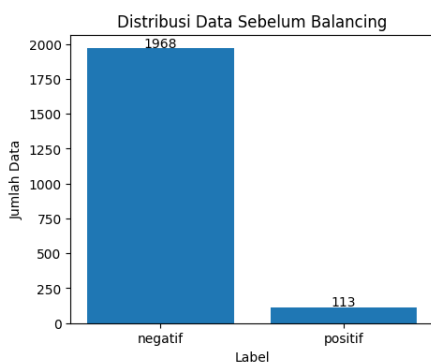
Dataset	Jumlah Data	Negatif	Positif
Training Set	2081	1968	113
Validation Set	446	422	24
Test Set	446	422	24

3.1.3 Tokenization dan Padding

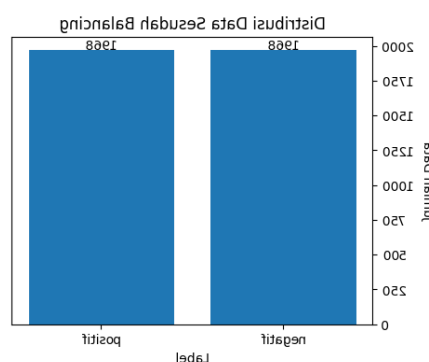
Proses tokenisasi dilakukan menggunakan *tokenizer* IndoBERTweet (*indolem/indobertweet-base-uncased*) dengan panjang maksimum sebanyak 50 token. Hasil tokenisasi menunjukkan adanya *subword tokenization* yang ditandai dengan token awalan “##” [9]. Selain itu, token [PAD] digunakan untuk menyamakan panjang input teks.

3.1.4 Balancing Data

Untuk mengatasi ketidakseimbangan data, penelitian ini menerapkan metode *Random Oversampling* pada *training set* dengan menambahkan data pada kelas minoritas hingga distribusi kedua kelas menjadi seimbang [13]. Proses ini bertujuan untuk mengurangi bias model terhadap kelas mayoritas dan meningkatkan kemampuan model dalam mengenali pola pada kelas minoritas sehingga performa klasifikasi menjadi lebih baik [14].



Gambar 3. Data Sebelum Balancing



Gambar 4. Data Sesudah Balancing

3.1.5 Pemodelan Analisis Sentimen Menggunakan IndoBERTweet

Pemodelan analisis sentimen dilakukan menggunakan model IndoBERTweet dengan dua kelas sentimen, yaitu positif dan negatif. Proses pelatihan menggunakan parameter *epoch* sebanyak 3, batch size 16, learning rate $2e-5$, dan *weight decay* 0,01. Model dilatih menggunakan *training set* dan dievaluasi menggunakan *validation set* pada setiap *epoch* untuk memantau performa model serta mengurangi risiko *overfitting* [15].

3.1.6 Evaluasi Model IndoBERTweet

Evaluasi model dilakukan menghitung metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

	precision	recall	f1-score	support
negatif	0.97	0.98	0.97	423
positif	0.52	0.46	0.49	24
accuracy			0.95	447
macro avg	0.75	0.72	0.73	447
weighted avg	0.95	0.95	0.95	447

Gambar 5. Pendekatan Pertama

	precision	recall	f1-score	support
negatif	0.99	0.99	0.99	423
positif	0.76	0.79	0.78	24
accuracy			0.98	447
macro avg	0.87	0.89	0.88	447
weighted avg	0.98	0.98	0.98	447

Gambar 6. Pendekatan Kedua

Hasil penelitian menunjukkan adanya perbedaan performa yang signifikan antara dua pendekatan dalam klasifikasi sentimen menggunakan model IndoBERTweet. Pendekatan pertama menggunakan *preprocessing* lengkap yang meliputi normalisasi, *stopword removal*, dan *stemming*, sedangkan pendekatan kedua mempertahankan teks dalam bentuk asli tanpa *preprocessing* agresif. Kedua pendekatan juga menerapkan *Random Oversampling* (ROS) untuk mengatasi ketidakseimbangan data.

Hasil evaluasi menunjukkan bahwa pendekatan pertama menghasilkan *accuracy* sebesar 95%, namun performa pada kelas minoritas masih rendah dengan *F1-score* sebesar 0.49, meskipun kelas mayoritas mencapai performa sangat tinggi. Hal ini mengindikasikan bahwa model masih cenderung bias terhadap kelas dominan. Kondisi ini diduga disebabkan oleh hilangnya informasi penting akibat *preprocessing* yang terlalu agresif, terutama dalam teks media sosial yang bersifat informal.

Sebaliknya, pendekatan kedua menunjukkan peningkatan yang signifikan dengan *accuracy* sebesar 98% dan *F1-score* kelas positif meningkat menjadi 0.78. Selain itu, nilai *macro F1-score* juga meningkat dari 0.73 menjadi 0.88, yang menunjukkan peningkatan keseimbangan performa antar kelas. Hasil ini menunjukkan bahwa mempertahankan teks dalam bentuk asli lebih sesuai untuk model IndoBERTweet yang telah dilatih pada data bahasa alami dan informal.

Meskipun demikian, metode ini memiliki kelemahan berupa potensi *overfitting* karena proses duplikasi sampel tanpa variasi data baru [16][17], yang dapat memengaruhi kemampuan generalisasi model. Secara keseluruhan, kombinasi *Random Oversampling* dan tanpa *preprocessing* agresif terbukti lebih efektif dalam meningkatkan performa klasifikasi, karena ROS membantu menyeimbangkan distribusi kelas, sementara penggunaan teks asli mempertahankan konteks linguistik yang penting bagi model.

3.1.7 Pemodelan Topik Menggunakan BERTopic

Pemodelan topik dilakukan menggunakan BERTopic dengan kombinasi *SentenceTransformer*, UMAP, HDBSCAN, dan c-TF-IDF. Proses ini menghasilkan 10 topik utama setelah dilakukan reduksi topik otomatis menggunakan fungsi *reduce_topics()*.

Tabel 3. Topic

Topic	Representation	Coherence
0	coretax, pajak, aplikasi, negara, pemerintah, penerimaan, sistem, buka, coretax coretax, indonesia	0.476175
1	error, error coretax, coretax error, error error, errornya, coretax errornya, bupot, sehari, pekerjaan, aplikasi error	0.444263
2	coretax coretax, coretax, birokrasi pusing, betulin, betulin coretax, coretax betulin, cabut coretax, coretax batin, bete	0.708003
3	djp, djp online, online, email, kpp, djp coretax, hp, coretax djp, no hp, data	0.833386
4	spt, ppn, pelaporan spt, spt ppn, pelaporan, posting, spt tahunan, posting spt, pph, ppn coretax	0.853327
5	npwp, nik, npwp coretax, registrasi npwp, npwp online, nomor, coretax npwp, aktivasi, online	0.782672
6	faktur, faktur pajak, pajak, upload faktur, pajak coretax, upload, fakturnya, pajak faktur, coretax upload, coretax faktur	0.861416
7	faktur pajak, faktur, pajak error, error, error coretax, pajak, coretax faktur, upload faktur, upload, keluaran	0.802024
8	website, website coretax, coretax website, mjb, website website, coretax mjb, mengikuti, kebanyakan, online website, webnya	0.342992
9	lag, coretax lag, lag coretax, sipd, lag lag, lag error, alasan gantimu, akunnnya, butuh restart, coretax kebangetan	0.686689

Hasil pemodelan menunjukkan bahwa topik yang muncul didominasi oleh isu teknis terkait implementasi Coretax, seperti *error* sistem, kendala *login* dan akses website, pelaporan SPT, faktur pajak, aktivasi akun, serta integrasi dengan DJP Online.

3.1.8 Evaluasi BERTopic

Evaluasi kualitas topik dilakukan menggunakan *coherence score* untuk mengukur keterkaitan semantik antar kata pada setiap topik. Hasil evaluasi menunjukkan rata-rata *coherence score* sebesar 0,64 yang mengindikasikan bahwa topik-topik yang dihasilkan cukup konsisten dan dapat diinterpretasikan dengan baik.

Beberapa topik memiliki *coherence* tinggi, seperti:

1. Topik 6 (Pengelolaan dan Upload Faktur Pajak) sebesar 0,861416
2. Topik 4 (Pelaporan SPT dan PPN di Coretax) sebesar 0.853327

Nilai tersebut menunjukkan bahwa kata-kata pada topik memiliki keterkaitan semantik yang kuat. Sementara itu, beberapa topik lain memiliki nilai *coherence* yang lebih rendah akibat penggunaan bahasa informal, singkatan, serta *typo* yang umum ditemukan pada data media sosial.

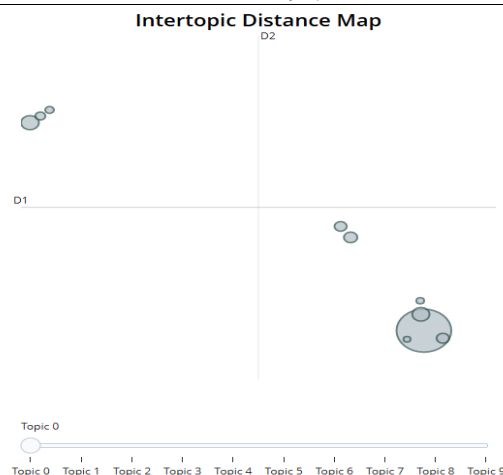
Penamaan topik dilakukan secara manual oleh peneliti dengan tujuan memberikan label yang representatif dan mudah dipahami [18].

Tabel 4. Penamaan Topik

No	Topik	Representation
0	Penurunan penerimaan APBN	coretax, pajak, aplikasi, negara, pemerintah, penerimaan, sistem, buka, coretax coretax, indonesia
1	Error Sistem Coretax	error, error coretax, coretax error, error error, errornya, coretax errornya, bupot, sehari, pekerjaan, aplikasi error
2	Keluhan dan Kritik terhadap Coretax	coretax coretax, coretax, birokrasi pusing, betulin, betulin coretax, coretax betulin, cabut coretax, coretax batin, bete
3	Akses Akun dan Integrasi DJP Online	djp, djp online, online, email, kpp, djp coretax, hp, coretax djp, no hp, data
4	Pelaporan SPT dan PPN di Coretax	spt, ppn, pelaporan spt, spt ppn, pelaporan, posting, spt tahunan, posting spt, pph, ppn coretax
5	Registrasi Wajib Pajak menggunakan NIK	npwp, nik, npwp coretax, registrasi npwp, npwp online, nomor, coretax npwp, aktivasi, online
6	Pengelolaan dan Upload Faktur Pajak	faktur, faktur pajak, pajak, upload faktur, pajak coretax, upload, fakturnya, pajak faktur, coretax upload, coretax faktur
7	Kendala Kendala Faktur Pajak dan Upload Dokumen	faktur pajak, faktur, pajak error, error, error coretax, pajak, coretax faktur, upload faktur, upload, keluaran
8	Akses Website Coretax	website, website coretax, coretax website, mjb, website website, coretax mjb, mengikuti, kebanyakan, online website, webnya
9	Kendala Website Coretax	lag, coretax lag, lag coretax, sipd, lag lag, lag error, alasan gantimu, akunya, butuh restart, coretax kebangetan

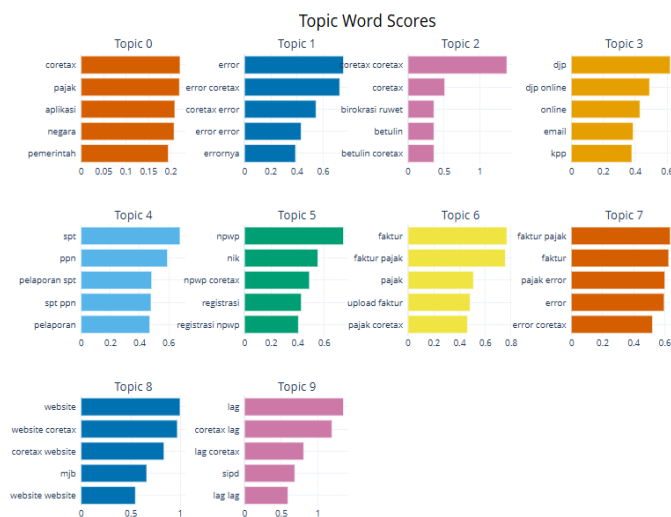
3.1.9 Visualisasi BERTopic

Visualisasi BERTopic dilakukan menggunakan *Topic Distance Map*, *Topic Word Score*, dan *Hierarchical Topics*. *Topic Distance Map* menunjukkan hubungan antar topik berdasarkan tingkat kemiripan pembahasan [19]. Berdasarkan visualisasi, terlihat bahwa beberapa topik berada dalam kelompok yang berdekatan, yang menandakan adanya keterkaitan tema antar topik. Selain itu, terdapat beberapa topik yang posisinya cukup jauh dari topik lainnya, sehingga dapat diinterpretasikan sebagai topik dengan pembahasan yang lebih spesifik atau berbeda secara signifikan.



Gambar 7. Topic Distance Map

Topic Word Score menampilkan kata-kata dengan bobot tertinggi pada setiap topik. setiap topik ditampilkan dalam bentuk diagram batang yang berisi kata-kata utama dengan nilai bobot tertinggi [20]. Kata-kata tersebut menggambarkan inti pembahasan dari masing-masing topik. Sebagai contoh, terdapat topik yang didominasi oleh kata seperti pajak, faktur, *website*, dan *login*, yang menunjukkan bahwa topik-topik tersebut berkaitan dengan layanan perpajakan digital, kendala akses sistem, maupun proses administrasi perpajakan.



Gambar 8. Topic Word Score

Sedangkan *Hierarchical Topics* membantu melihat struktur hubungan antar topik dalam bentuk dendrogram [20]. Pada *dendrogram*, topik yang memiliki kemiripan tinggi akan bergabung terlebih dahulu pada jarak yang lebih dekat. Sebaliknya, topik yang berbeda akan bergabung pada jarak yang lebih jauh. Berdasarkan visualisasi, terlihat adanya beberapa kelompok topik yang saling berkaitan, seperti topik mengenai akses website, login, dan aplikasi perpajakan yang berada dalam satu kelompok besar. Selain

itu, terdapat pula kelompok topik lain yang berkaitan dengan faktur pajak, pelaporan SPT, dan kendala sistem.



Gambar 9. Hierarchical Topic

4. SIMPULAN

Data penelitian yang digunakan sebanyak 3.253 tweet yang dikumpulkan dari media sosial X, dengan distribusi label terdiri atas 2.794 data berlabel negatif dan 459 data berlabel positif. Ketidakseimbangan jumlah data pada masing-masing kelas menunjukkan bahwa respons publik terhadap implementasi Coretax cenderung didominasi oleh sentimen negatif. Hasil penelitian ini berhasil mengimplementasikan IndoBERTweet dan BERTopic untuk analisis sentimen dan pemodelan topik terhadap tweet terkait implementasi Coretax. Model IndoBERTweet menunjukkan performa yang baik dalam klasifikasi sentimen dengan hasil terbaik pada pendekatan tanpa preprocessing agresif, memperoleh accuracy sebesar 92% dan F1-score sebesar 82%. Selain itu, BERTopic berhasil mengidentifikasi 14 topik utama yang sebagian besar membahas isu gangguan teknis, seperti error sistem, kendala login, pelaporan SPT, dan pengelolaan faktur pajak. Dengan nilai rata-rata coherence sebesar 0,54 menunjukkan bahwa topik yang dihasilkan cukup konsisten dan dapat diinterpretasikan dengan baik.

REFERENCES

- [1] F. Faradina, "Coretax DJP: 1 Aplikasi 7 Manfaat," Direktorat Jendral Pajak. Accessed: Sep. 25, 2025. [Online]. Available: [https://www.pajak.go.id/id/artikel/coretax-djp-1-aplikasi-7-manfaat#:~:text=Tujuan utama dari pembangunan Coretax,hingga pemeriksaan dan penagihan pajak](https://www.pajak.go.id/id/artikel/coretax-djp-1-aplikasi-7-manfaat#:~:text=Tujuan%20utama%20dari%20pembangunan%20Coretax,hingga%20pemeriksaan%20dan%20penagihan%20pajak)
- [2] S. P. Tamba, "Dari Layar Lebar Menuju Era Perpajakan Digital," Direktorat Jendral Pajak. Accessed: Sep. 25, 2025. [Online]. Available: <https://www.pajak.go.id/id/artikel/dari-layar-lebar-menuju-era-perpajakan-digital>
- [3] B. Liu, *Sentiment Analysis*. Cambridge University Press, 2020. doi: 10.1017/9781108639286.
- [4] B. E. S. Dewi, "Model Analisis Sentimen Pada Kendaraan Listrik Menggunakan Algoritma IndoBERTweet Ddan IndoBER," vol. 19, no. 1, pp. 169–179, 2025, doi: <https://doi.org/10.35457/w5r3g517>.
- [5] A. S. Rizkia, W. Wufron, and F. F. Roji, "Analisis Sentimen Coretax: Perbandingan Pelabelan Data Manual, Transformers-Based, dan Lexicon-Based pada Performa IndoBERT," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 3, Jul. 2025, doi: 10.57152/malcom.v5i3.2151.

-
- [6] Nursyahrina, S. Defit, and R. Sovia, "Metode BERTopic dan LDA untuk Analisis Tren Penelitian Bidang Ilmu Komputer," *J. KomtekInfo*, vol. 11, no. 4, pp. 332–341, 2024, doi: 10.35134/komtekinfo.v11i4.580.
 - [7] M. W. A. Putra, Susanti, Erlin, and Herwin, "Analisis Sentimen Dompot Elektronik Pada Media Sosial Twitter Menggunakan Naïve Bayes Classifier Pengguna (dalam Porsen)," *IT J. Res. Dev.*, vol. 5, no. 1, pp. 72–86, 2020.
 - [8] I. Nura Nugraha and E. Utami, "Evaluation of Creative Economy and Tourism Industry Trends based on LDA Analysis with BERTopic," *Digit. Zo. J. Teknol. Inf. dan Komun.*, vol. 15, no. 2, pp. 182–195, 2024, doi: 10.31849/digitalzone.v15i2.23796.
 - [9] F. Koto, J. H. Lau, and T. Baldwin, "INDOBERTWEET: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization," *EMNLP 2021 - 2021 Conf. Empir. Methods Nat. Lang. Process. Proc.*, pp. 10660–10668, 2021, doi: 10.18653/v1/2021.emnlp-main.833.
 - [10] P. Kherwa and P. Bansal, "EAI Endorsed Transactions on Scalable Information Systems Topic Modeling : A Comprehensive Review," vol. 7, no. 24, pp. 1–16, doi: <https://doi.org/10.4108/eai.13-7-2018.159623>.
 - [11] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022, [Online]. Available: <http://arxiv.org/abs/2203.05794>
 - [12] D. H. Seol, J. E. Choi, C. Y. Kim, and S. J. Hong, "Alleviating Class-Imbalance Data of Semiconductor Equipment Anomaly Detection Study," *Electronics*, vol. 12, no. 3, p. 585, Jan. 2023, doi: 10.3390/electronics12030585.
 - [13] S. Diantika, "Penerapan Teknik Random Oversampling Untuk Mengatasi Imbalance Class Dalam Klasifikasi Website Phishing Menggunakan Algoritma Lightgbm," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 7, no. 1, pp. 19–25, 2023, doi: 10.36040/jati.v7i1.6006.
 - [14] R. Aryanti, T. Misriati, and R. Hidayat, "Klasifikasi Risiko Kesehatan Ibu Hamil Menggunakan Random Oversampling Untuk Mengatasi Ketidakseimbangan Data," *KLIK Kaji. Ilm. Inform. dan Komput.*, vol. 3, no. 5, pp. 409–416, 2023, doi: <https://doi.org/10.30865/klik.v3i5.728>.
 - [15] M. N. Zaidan, Y. Sibaroni, and S. S. Prasetyowati, "LEARNING RATE AND EPOCH OPTIMIZATION IN THE FINE-TUNING PROCESS FOR INDOBERT'S PERFORMANCE ON SENTIMENT ANALYSIS OF MYTELKOMSEL APP REVIEWS," *J. Tek. Inform.*, vol. 5, no. 5, pp. 1443–1450, Oct. 2024, doi: 10.52436/1.jutif.2024.5.5.2396.
 - [16] Y. Yang, H. A. Khorshidi, and U. Aickelin, "A review on over-sampling techniques in classification of multi-class imbalanced datasets: insights for medical problems," *Front. Digit. Heal.*, vol. 6, no. July, 2024, doi: 10.3389/fdgth.2024.1430245.
 - [17] C. Yang, E. A. Fridgerisson, J. A. Kors, J. M. Reys, and P. R. Rijnbeek, "Impact of random oversampling and random undersampling on the performance of prediction models developed using observational health data," *J. Big Data*, vol. 11, no. 1, 2024, doi: 10.1186/s40537-023-00857-7.
 - [18] S. J. Weston, I. Shryock, R. Light, and P. A. Fisher, "Selecting the Number and Labels of Topics in Topic Modeling: A Tutorial," *Adv. Methods Pract. Psychol. Sci.*, vol. 6, no. 2, 2023, doi: 10.1177/25152459231160105.
 - [19] C.- Crisis, S. Avasthi, R. Chauhan, and D. P. Acharjya, *International Conference on Innovative Computing and Communications*, vol. 1387, no. January. in *Advances in Intelligent Systems and Computing*, vol. 1387. Singapore: Springer Singapore, 2022. doi: 10.1007/978-981-16-2594-7.
 - [20] D. Aryani, I. Lucia Kharisma, A. Sujjada, and K. Kamdan, "Topic Modeling of the 2024 Election Using the BERTopic Method on Detik.com News Articles," *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 9, no. 2, pp. 171–180, 2024, doi: 10.25139/inform.v9i2.8429.